



(19) Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) EP 1 424 395 A1

(12) **EUROPEAN PATENT APPLICATION**  
published in accordance with Art. 158(3) EPC

(43) Date of publication:  
02.06.2004 Bulletin 2004/23

(51) Int Cl.7: **C12N 15/52, C12N 9/00,**  
**C07B 61/00, G01N 33/48**  
**// (G06F17/30, 17:50)**

(21) Application number: 02794635.9

(22) Date of filing: 11.01.2002

(86) International application number:  
**PCT/JP2002/000118**

(87) International publication number:  
**WO 2003/014354 (20.02.2003 Gazette 2003/08)**

(84) Designated Contracting States:  
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU**  
**MC NL PT SE TR**

• **SAKAMOTO, Kensaku**  
Adachi-ku, Tokyo 120-0044 (JP)  
• **HIRAO, Ichiro**  
Asaka-shi, Saitama 351-0036 (JP)  
• **YOKOYAMA, Shigeyuki**  
Bunkyo-ku, Tokyo 113-0023 (JP)

(30) Priority: 01.08.2001 JP 2001234135

(74) Representative: **Cresswell, Thomas Anthony**  
**J.A. KEMP & CO.**  
**14 South Square**  
**Gray's Inn**  
**London WC1R 5JJ (GB)**

(71) Applicant: **Japan Science and Technology Agency**  
Kawaguchi-shi, Saitama 332-0012 (JP)

(72) Inventors:  
• **KIGA, Daisuke**  
Wako-shi, Saitama 351-0112 (JP)

(54) **TYROSYL-tRNA SYNTHASE VARIANTS**

(57) Tyrosyl-tRNA synthetases having a modified amino acid sequence whereby unnatural amino acids can be more efficiently incorporated than original natural amino acids. More specifically, tyrosyl-tRNA synthetases having been modified at the amino acid(s) at the 37- and/or 195- positions. A method of modifying tyrosyl-

tRNA synthetase characterized in that the position of an amino acid to be modified is determined based on the 3-D structure of a tyrosyl-AMP and of tyrosyl synthetase complex and an amino acid to which this enzyme binds.

**Description****Technical field**

5 [0001] The present invention relates to tyrosyl-tRNA synthetase of which amino acid sequence is modified whereby capable of incorporating unnatural tyrosine analogue substituted at position 3 more efficiently than original natural tyrosine by modifying the amino acid sequence of the tyrosyl-tRNA synthetase, and preferably to tyrosyl-tRNA synthetase variants where two or more amino acids are modified. The tyrosyl-tRNA synthetase of the invention of which amino acid sequence is modified can incorporate unnatural tyrosine analogue substituted at position 3, e.g., 3-iodo-  
10 tyrosine more efficiently than original natural tyrosine, and more intensively utilize it than tyrosine which is a substrate of wild-type of tyrosyl-tRNA synthetase. Therefore, the use thereof can produce proteins (alloproteins) containing, unnatural amino acids selectively and site-specifically at a high efficiency.

**Background art**

15

[0002] Variant proteins have been produced where a certain amino acid residue of a protein is substituted with an amino acid other than 20 canonical amino acids (referred to as an unnatural amino acid hereinafter) involved in typical protein synthesis. It has been advocated that the protein which contains unnatural amino acids is referred to as the alloprotein (Koide et al., Proc. Natl. Acad. Sci. USA, 85:6237-41, 1988). It makes finer and systematic modification possible that the certain residue can be also replaced by the unnatural amino acid, compared to the cases where the replacements occur among 20 types of natural amino acids. Also, the amino acid with characteristic fluorescent property, the amino acid of which structure can be optically controlled, the amino acid with a reaction group applicable as a optic crosslinker, and the like have been introduced into the proteins.

20

[0003] There are some techniques to produce the alloproteins. Koide et al. made *Escherichia coli* incorporate an unnatural amino acid added to medium, which was used instead of one certain type of the canonical amino acids to produce the alloprotein. However, only 20 types together with canonical amino acids and unnatural amino acids can be used in this technique (Koide et al.; Proc. Natl. Acad. Sci. USA, 85:6237-41, 1988).

25

[0004] Alternatively, the alloprotein is produced by adding the suppressor tRNA which has been aminoacylated beforehand in a separate system to a cell-free translation system (Noren et al., Science, 244:182-8, 1989). The disadvantage of this method includes a necessity to prepare aminoacyl-tRNA in a large amount.

30

[0005] In order to prepare the protein comprising 21 types of amino acids including the unnatural amino acid in the larger amount, it is necessary to construct an artificial genetic code system in which the tRNA attaching the unnatural amino acid is aminoacylated by its cognate aminoacyl-tRNA synthetase (aaRS) in the system where the translation reaction is conducted. The aminoacyl-tRNA synthetase is an enzyme which specifically attaches the amino acid to the tRNA, and 20 types occur corresponding to respective 20 types of canonical amino acids for each biological species excluding some exceptions. In a cell, these enzymes determine the type of the amino acid assigned to the genetic code where basically one type of such an aaRS exists for every amino acid. For instance, tyrosyl-tRNA synthetase (TyrRS) which is one of aminoacyl-tRNA synthetase (aaRS) discriminates tRNA for tyrosine from the other tRNA for the other amino acids, and makes it attach to only tyrosine but not the other amino acids.

35

[0006] In the meanwhile, Wang et al. expressed TyrRS variants derived from *Methanococcus janasii* modified so as to attach O-methyltyrosine specifically and amber suppressor tRNA engineered by the modification of tyrosine tRNA derived from the same organism, in *E. coli* (Wang et al., Science, 292:498-500, 2001). This TyrRS and tyrosine tRNA from *Methanococcus janasii* do not react with tRNA and aaRS from *E. coli*, respectively. Thus, it has been reported that O-methyltyrosine is incorporated specifically corresponding to an amber codon in this study.

40

[0007] In order to construct such an artificial genetic code system, it is critical that a pair of aaRS and tRNA which does not react with aaRS from the host is found and that the aaRS variant which reacts specifically with the unnatural amino acid is developed. As the aaRS which reacts specifically with the unnatural amino acid, only TyrRS modification specific for O-methyltyrosine mentioned above has been known so far.

45

[0008] As tyrosine analogues substituted at position 3, there are DOPA involved in intercellular signal transduction, 3-iodotyrosine capable of becoming a target site of site-specific labeling in the protein, and the like. These have been known as the unnatural amino acids which exert physiological activity. Therefore, it is desired to obtain TyrRS specific for the tyrosine analogues substituted at position 3. No TyrRS variant has been known so far, which incorporates such tyrosine analogues substituted at position 3 more efficiently than tyrosine. The TyrRS variant was reported at an academic meeting, which incorporates the tyrosine analogues substituted at position 3 more efficiently than the wild-type TyrRS. However, since this variant incorporates tyrosine and the tyrosine analogue at a similar efficiency, it is inappropriate for inserting only the unnatural amino acid in the certain site of the protein.

Disclosure of the invention

[0009] The present invention provides a novel tyrosyl-tRNA synthetase variant which can incorporate preferentially an unnatural amino acid, more particularly tyrosine analogue having the substitution at position 3 of phenyl group of the tyrosine. More particularly, provided is a novel tyrosyl-tRNA synthetase variant wherein the unnatural amino acid is 3-halotyrosine such as 3-iodotyrosine.

[0010] Also, the invention provides a novel method for producing protein containing the unnatural amino acid (allo-protein) by a protein production system using the novel tyrosyl-tRNA synthetases which can preferentially incorporate the unnatural tyrosine analogue substituted at position 3.

[0011] Further, the invention provides a method for designing the tyrosyl-tRNA synthetases which can preferentially and efficiently incorporate the unnatural tyrosine analogues substituted at position 3, and cells transformed by a gene encoding the tyrosyl-tRNA synthetase designed in such a method and production of the proteins containing the unnatural amino acid in the cells, as well as a production means of the proteins including the unnatural amino acid in a cell-free translation system to which the designed aminoacyl-tRNA synthetase is added.

Brief Description of the Drawings**[0012]**

Figure 1 is a photo, substitute for a drawing, showing the results of aminoacylation activity where the aminoacyl-tRNA synthetase variants of the invention were assayed by the same method as acidic polyacrylamide gel electrophoresis.

Figure 2 shows amber suppression in a wheat germ cell-free translation system. Translation products labeled with <sup>14</sup>C-leucine were analyzed by SDS-PAGE. The amber suppression was observed only when both TyrRS-VC variant and suppressor tRNA were present. A symbol, amb is mRNA in which the 32nd codon was replaced by the amber codon, and sup tRNA is the amber suppressor tRNA derived from tRNA<sup>Tyr</sup> from *E. coli*.

Figure 3 shows charts of liquid chromatography in LC-MS analysis of translation products in the wheat germ cell-free translation system in which iodotyrosine, TyrRS-VC variant specific for iodotyrosine and the suppressor tRNA were included.

K3 and K9 are lysylendopeptidase digested fragments of Ras protein, respectively. They were found to be Ser 17-Lys 42 fragment and Thr 148-Lys 167 fragment by mass spectrometry. The upper and lower panels show the analysis of the translation products of mRNA comprising no amber codon and mRNA in which 32nd codon is the amber codon, respectively. A fragment (K3-IY32) having a mass corresponding to the fragment where one tyrosine residue of the K3 fragment was replaced by iodotyrosine was found in the analysis of the product of the amber suppression (lower panel).

Best Mode for Carrying Out the Invention

[0013] The present invention relates to tyrosyl-tRNA synthetases having a modified amino acid sequences whereby capable of incorporating unnatural tyrosine analogues substituted at position 3 more efficiently than an original natural tyrosine. More particularly, the invention relates to the tyrosyl-tRNA synthetases capable of incorporating amino acids substituted at position 3 of tyrosine which are unnatural type of amino acids efficiently and selectively, by modifying two or more sites of amino acids of the amino acid sequence of the tyrosyl-tRNA synthetase at residues presumed to recognize the target tyrosine with the other amino acids on the basis of its tertiary structure.

[0014] Also, the invention relates to a method for modifying the amino acid sequence of the tyrosyl-tRNA synthetase. This modification is characterized in that amino acids at the site where the tyrosyl-tRNA synthetase binds to tyrosine or tyrosyl-AMP are determined on the basis of its three dimensional structure and the amino acids are modified with the other amino acids, preferably two or more amino acids are modified with the other amino acids, when producing the tyrosyl-tRNA synthetase having the modified amino acid sequence whereby capable of incorporating the unnatural tyrosine analogue substituted at position 3 more efficiently than the original natural tyrosine by modifying the amino acid sequence of the tyrosyl-tRNA synthetase.

[0015] Further, the invention relates to a method for producing proteins containing the unnatural amino acids using the tyrosyl-tRNA synthetases having the modified amino acid sequence of the invention. More particularly, the invention relates to a method for producing the proteins containing the tyrosine analogues substituted at position 3 of tyrosine which are the unnatural type amino acids using the tyrosyl-tRNA synthetases having the modified amino acid sequence.

[0016] Also, the invention relates to cells transformed by a gene encoding such a tyrosyl-tRNA synthetase having the modified amino acid sequence, and production of proteins comprising the unnatural amino acids in such cells as

well as a production means of the proteins comprising the unnatural amino acids in a cell-free translation system to which the designed aminoacyl-tRNA synthetase is added.

[0017] The present inventors have studied on methods capable of producing alloproteins efficiently, selectively and in a large scale, and developed the aminoacyl-tRNA synthetases capable of incorporating the unnatural type amino acid preferentially and efficiently therefor. Tyrosyl-tRNA synthetase (TyrRS) is one type of the aminoacyl-tRNA synthetases, and the present inventors focused on the tyrosine analogue recognizing tyrosyl-tRNA synthetase (TyrRS) variants which recognize tyrosine analogues since the tyrosine analogues substituted at position 3 per se have various types of physiological activity. Also, it has been known that TyrRS and tyrosine tRNA derived from *E. coli* do not react with tyrosine tRNA and TyrRS from eukaryotic organisms, respectively. Thus, a TyrRS variant specific for the tyrosine analogues substituted at position 3 is also promising as a material to construct an artificial genetic code system for the site-specific incorporation of the unnatural amino acid into a protein.

[0018] Therefore, the following illustration is carried out citing the tyrosyl-tRNA synthetase (TyrRS) as a specific example of the aminoacyl-tRNA synthetases.

[0019] The present inventors studied on the improvement of the amino acids at a substrate binding site of the enzyme to obtain TyrRS variants specific for the tyrosine analogues substituted at position 3. The amino acid sequence of TyrRS derived from *E. coli* has been already known and it is shown at the sequence No. 1 in the sequence listing. Also the amino acid sequence is shown below using the single character code of amino acids.

20	MASSNLIKQL QERGLVAQVT DEEALALAERIA	30
25	QGPIAL <u>Y</u> CGF DPTADSLHLG HLVPLLCLKR	60
30	FQQAGHKPVA LVGGATGLIG DPSFKAAERK	90
35	LNTEETVQEW VDKIRKQVAP FLDLDCGENS	120
40	AIAANNYDW <u>F</u> GNMNVLTFLR DIGKHFSVNQ	150
45	MINKEAVKQR LNREDQGISF TEFSYNLLQG	180
50	YDFACLNKQY GVVL <u>Q</u> IGGSD QWGNITSGID	210
55	LTRRLHQNQV FGLTVPLIK ADGTKFGKTE	240
60	GGAVWLDPKK TSPYKFYQFW INTADADVYR	270
65	FLKFFTMSI EEINALEEEED KNSGKAPRAQ	300
70	YVLAEQVTRL VHGE EGLQAA KRITECLFSG	330
75	SLSALSEADF EQLAQDGVP M VEMEKGADLM	360
80	QALVDSELQP SRGQARKTIA SNAITINGEK	390
85	QSDPEYFFKE EDRLFGRFTL LRRGKKNYCL	420
90	ICWK	424

[0020] The present inventors studied and determined which residues on the enzyme should be replaced as follows.

Three dimensional structure of *Bacillus stearothermophilus* TyrRSV-tyrosyl-AMP complex has been already reported (Brick et al., J. Mol. Bio., 208:83, 1988), and the residues where the tyrosine analogues substituted at position 3 should be incorporated were presumed on the basis of this three dimensional structure. Although this is a relationship of the TyrRS-tyrosyl-AMP complex in *Bacillus stearothermophilus*, the present inventors supposed that it is possible to figure similarly for TyrRS derived from the other organisms because of highly conservative TyrRS in biological species.

5 [0021] Based on this three dimensional structure model, the present inventors found the amino acids at positions 37 and 195 as the residues where the replacements should be introduced into TyrRS of *E. coli* so as to incorporate the tyrosine analogues substituted at position 3. The amino acid at position 37 is tyrosine (Y) and the amino acid at position 10 195 is glutamine (Q) in the wild-type TyrRS. They are the underlined positions in the amino acid sequence described above.

[0022] First, an expression vector, pET-YRS for TyrRS was made by preparation of the fragment amplified by PCR using the following primers (1) and (2) and using a plasmid including TyrRS gene of *E. coli* as a template, and digestion of the fragment with NdeI and HindIII, followed by the insertion into a NdeI-HindIII site of pET26b.

15

Primer (1): 5'-GGAATTCCATATGGCAAGCAGTAACTTGATTAAACAATT  
GCAAG-3'

20

Primer (2): 5'-GCCGAAGCTTGTCGACTTCCAGCAAATCAGACAGTAATTCT  
TTTACCG-3'

25

[0023] Next, the method for the site-directed mutagenesis of the amino acids at the residues 37 and/or 195 is illustrated.

30 [0024] First, a single amino acid substitution was made where the amino acid at either position of 37 or 195 was substituted. The primers (3) to (8) used for making the DNA sequences encoding the substitutions of one amino acid at positions 37 and 195 are shown below.

35

Primer (3): 5'-AGGATCGAAGCCGCAAGCGAGCGCGATCGGGCTTGCGCC-3'

40

Primer (4): 5'-AGGATCGAAGCCGCANNNGAGCGCGATCGGGCTTGCGCC-3'

45

[0025] M denotes C or A, and N denotes A, C, G or T.

50

Primer (5): 5'-ACGGTGTGGTGCTGTCTATTGGTGGTTCTGACC-3'

55

Primer (6): 5'-ACGGTGTGGTGCTGGCAATTGGTGGTTCTGACC-3'

50

Primer (7): 5'-ACGGTGTGGTGCTGAACATTGGTGGTTCTGACC-3'

55

Primer (8): 5'-ACGGTGTGGTGCTGTGCATTGGTGGTTCTGACC-3'

[0026] Next, a double amino acid substitution was made where both amino acids at the position 37 and 195 were modified.

[0027] The DNA sequence encoding the double amino acid substitution was made by an overlap extension method using the primers from the plasmids encoding the respective single amino acid substitutions at the residues 37 and 195 made in the above step, and inserted into an NdeI-BamHI site of pET-YRS. The overlap extension method was carried out by purifying two fragments obtained by the amplification using a pair of the primers (1) and (10) and a pair of the primers (9) and (11), and subsequently amplifying by PCR using the primers (1) and (9) therewith.

10 Primer (1): 5'-GGAATTCCATATGGCAAGCAGTAACITGATTAAACAATT  
GCAAG-3'

15 Primer (10): 5'-GATCATCTGGTTAACGGAGAAGTGTGTTGCC-3'

20 Primer (9): 5'-TTCTTCGGATCCAACCAACCGAGACTGCGCCGCCCTC-3'

25 Primer (11): 5'-GACCTTCCTGCGCGATATTGGCAAAC-3'

[0028] Modifications of the amino acids obtained in this way and their symbols are shown in the following Table 1.

Table 1.

Symbols of wild-types and modifications of amino acids and types of amino acids			
	Symbol	Position 37	Position 195
	Wild-type	Tyrosine (Y)	Glutamine (Q)
35	37A	Alanine (A)	Glutamine (Q)
	37V	Valine (V)	Glutamine (Q)
	37I	Isoleucine (I)	Glutamine (Q)
	37L	Leucine (L)	Glutamine (Q)
	195S	Tyrosine (Y)	Serine (S)
40	195C	Tyrosine (Y)	Cysteine (C)
	195N	Tyrosine (Y)	Asparagine (N)
	195A	Tyrosine (Y)	Alanine (A)
	AS	Alanine (A)	Serine (S)
	VS	Valine (V)	Serine (S)
45	IS	Isoleucine (I)	Serine (S)
	LS	Leucine (L)	Serine (S)
	AC	Alanine (A)	Cysteine (C)
	VC	Valine (V)	Cysteine (C)
50	IC	Isoleucine (I)	Cysteine (C)
	LC	Leucine (L)	Cysteine (C)
	AN	Alanine (A)	Asparagine (N)
	VN	Valine (V)	Asparagine (N)
	IN	Isoleucine (I)	Asparagine (N)
55	LN	Leucine (L)	Asparagine (N)
	AA	Alanine (A)	Alanine (A)
	VA	Valine (V)	Alanine (A)

Table 1. (continued)

Symbols of wild-types and modifications of amino acids and types of amino acids		
Symbol	Position 37	Position 195
IA	Isoleucine (I)	Alanine (A)
LA	Leucine (L)	Alanine (A)

[0029] Each complete variant DNA fragment obtained at the above step was inserted into an original position in the plasmid pETYRS. *E. coli* BLR(DE3) was transformed with the plasmid including the wild-type and variant-type tyrS genes by the transformation method according to Hanahan's method (Hanahan, D., J. Mol. Bio., 166:557-580). The wild-type and variant-type TyrRS were expressed by isolating and culturing the transformants having respective plasmids. The wild-type and variant-type TyrRS proteins were purified from the *E. coli* crude extracts.

[0030] Assays of substrate specificity were carried out for the resultant modifications. The following three assay methods were used for the assays.

[0031] In the first method, Lloyd's method (Nucleic Acids Research, 23:2886-2892, 1995) where inorganic phosphate produced by pyrophosphatase-driven hydrolysis of pyrophosphate which was one of the reaction products of the aminoacylation reactions was quantified was simplified, and the inorganic phosphate was detected using Biomol green (Funakoshi) to carry out the measurement of the aminoacylation reaction. The results are shown in Tables 2 and 3.

[0032] In the second method, Lloyd's method (Nucleic Acids Research, 23:2886-2892, 1995) was carried out without alteration. The results are shown in Table 4.

[0033] In the third method, levels of aminoacylation activity in the presence of tyrosine or 3-iodotyrosine for the full-length tRNA were assayed by the same method as acidic polyacrylamide gel electrophoresis described by Wolfson et al. (Wolfson, A.D. et al., RNA, 4:1019-1023, 1998). The result is shown in Figure 1 as a photo, substitute for a drawing.

25

Table 2.

Aminoacylation activity of wild-type and single amino acid variant enzymes		
	Tyrosine	3-iodotyrosine
Wild-type	***	-
37A	***	***
37V	***	***
37I	**	***
37L	***	***
195S	***	*
195C	***	*
195N	***	*
195A	***	*

40

[0034] Marks in the table denote that the activity is high in the order of \*\*\* > \*\* > \* > -.

[0035] Each mark shows that the aminoacylation reaction was detected by a microplate reader in the reaction at enzyme concentrations of 50 nM, 500 nM and 5  $\mu$ M at 37°C for 4 min. Amino acid concentrations were 200  $\mu$ M and 1 mM for tyrosine and 3-iodotyrosine, respectively.

45

Table 3.

Aminoacylation activity of wild-type, 37V and double variant enzymes		
	Tyrosine	3-iodotyrosine
Wild-type	***	-
37V	***	***
AS	*	*
VS	**	**
IS	-	-
LS	-	-
AG	*	**

Table 3. (continued)

Aminoacylation activity of wild-type, 37V and double variant enzymes			
	Tyrosine	3-iodotyrosine	
5	VC	*	**
	IC	-	-
	LC	*	*
10	AN	-	-
	VN	-	**
	IN	-	-
	LN	*	-
15	AA	-	-
	VA	*	**
	IA	*	*
	LA	*	-

[0036] Marks in the table denote that the activity is high in the order of \*\*\*, > \*\* > \* > -.

[0037] Each mark shows that the reaction could be detected strongly at 25 nM, strongly at 250 nM and weakly at 250 nM of enzyme concentrations in the amino acylation reaction at 37°C for 4 min. Amino acid concentrations were 200 μM and 1 mM for tyrosine and 3-iodotyrosine, respectively.

[0038] Table 2 lists the levels of aminoacylation activity of the wild-type enzyme and 8 variant enzymes with a single amino acid mutation, and Table 3 lists the levels of aminoacylation activity of the wild-type, 37V enzymes and 16 variant enzymes with double mutations. Concerning the enzyme variants obtained by substitutions of Tyr 37 and Gln 195 with the other amino acids, the levels of amino acylation activity were substantially increased against the tyrosine analogue substituted at position 3 (3-iodotyrosine in this experiment). On the other hand, relatively, the levels of aminoacylation activity of these proteins specific for tyrosine are decreased. Therefore, the enzyme variants with relatively higher ratio of tyrosine analogue to tyrosine than that of the wild-type emerged.

Table 4.

Aminoacylation activity of enzymes with double mutations Initial velocity of release of pyrophosphates (nM PPi/min/nM enzyme) at 37°C			
	Tyrosine	3-iodotyrosine	
35	Wild-type	400	<0.05
	37V	140	73
	37A	51	12
40	AC	1.0	3.1
	VC	0.9	9.5
	VS	19	18
	VN	2.1	6.9
45	VA	3.7	6.6

[0039] Table 4 lists the levels of aminoacylation activity of the enzymes with double mutations determined by Lloyd's method described in Nucleic Acids Research, 23:2886-2892, 1995. Four enzymes with double mutations combine the advantage of variants at both Tyr 37 and Gln 195. For instance, the relative tyrosine analogue to tyrosine ratio of VC is more than 10, and the ratio of VC is higher than that of 37V, the enzyme with a single mutation, due to further introduction of the replacement at Gln 195. VN, VS and AC are inferior to VC in the relative tyrosine analogue to tyrosine ratio, but have the higher relative tyrosine analogue to tyrosine ratio than those of the substitution with a single amino acid.

[0040] Also from the result shown in Table 4, it was concluded that the combination of the variants at positions 37 and 195 dramatically altered the substrate specificity of the enzyme and that the level of aminoacylation activity for the tyrosine analogue substituted at position 3 could be increased. It is also found that VC, VN and AC have the most notable property and applicability among the enzyme variants.

[0041] Figure 1 is a photo, substitute for a drawing, showing the aminoacylation activity of the enzymes with double

mutations. The assay was carried out for the VC having the highest tyrosine analogue to tyrosine ratio by acidic PAGE where an actual attachment of an amino acid to tRNA could be examined. As a result, it was found that VC had a high specificity for the tyrosine analogue.

[0042] Production of the protein containing iodotyrosine was attempted by the reaction by a batch method in a wheat germ cell-free protein synthesis system using the resultant VC and an amber suppressor tRNA derived from tyrosine tRNA of *E. coli*. A short chain type of c-H-Ras added His-tag at the C-terminus via a spacer sequence (Milburn et al., Science 247:939-45, 1990) was used as the target protein. In translation where the template was mRNA made from a construction where the 32nd codon of this protein was replaced by the amber codon, the production of the protein in full length was dependent on coexistence of both TyrRS and the suppressor tRNA (Figure 2). This proves that the suppressor tRNA used is scarcely aminoacylated by the endogenous aminoacyl-tRNA synthetase in the wheat germs. A synthetic quantity by the amber suppression was 30 to 40% of that of a control with no amber codon. In order to increase the synthetic quantity of the protein, the reaction by the suppression was carried out for 3 days using a dialysis method. In this case, it was found that the quantity of the production was 0.1 mg or more in terms of the quantity per ml of reaction solution and that the protein could be synthesized sufficiently in quantity to perform various experiments using proteins.

[0043] It was examined by mass spectrometry that iodotyrosine was actually inserted at the position corresponding to the amber codon of the peptide chain. Ras protein purified using nickel beads was applied on SDS-PAGE. A band corresponding to a mobility of Ras among stained product bands was cut out and treated with peptidase. A resultant fragment was analyzed by LC-MS (liquid chromatography - mass spectrometry). The 32nd codon of Ras protein used as a reporter protein is the codon for tyrosine. For mRNA where this codon was modified with the amber codon and for intact mRNA, the translation reaction was carried out where iodotyrosine, VC variant and the suppressor tRNA were all added to the translation system in the reaction solution. Among the fragments obtained by digesting the translation products, the fragment including the 32nd amino acid from N-terminus (Ser 17 to Lys 42, referred to as K3 fragment) was focused and analyzed. As a result, the majority of the K3 fragments produced by the amber suppression had the mass corresponding one where one residue of tyrosine was replaced by iodotyrosine, and K3 fragments not including iodotyrosine was minor. No K3 fragment was detected where the other amino acid was accidentally inserted at the site corresponding to the amber codon. The ratio of iodotyrosine to tyrosine inserted to the amber codon was examined by absorbance of chromatogram. Consequently, it was found that 95% or more of K3 fragments encompassed iodotyrosine by the amber suppression whereas the ratio of tyrosine insertion at the amber codon was 5% or less. On the other hand, no fragment having the mass corresponding to the case encompassing iodotyrosine in place of tyrosine was detected in the analysis of the translation products from mRNA which contained no amber codon. This confirmed that the canonical codon for tyrosine was not contaminated with iodotyrosine.

[0044] The method of the invention was illustrated by the use of Ras protein as an example, but it is obvious that the method of the invention is not limited to Ras protein and can be similarly applied for the other proteins.

[0045] As described above, the tyrosyl-tRNA synthetase was illustrated by citing TyrRS of *E. coli* as the specific example. It has been shown by the specific example that various enzyme variants having tyrosylation activity where the relative ratio of the unnatural amino acid, tyrosine analogue substituted at position 3 to tyrosine, is high can be found by converting tyrosine at position 37 and glutamine at position 195 to the other amino acids in the wild-type TyrRS of *E. coli*. Also, such enzyme variants made it possible to apply for the production of the protein with site-specific incorporation of the tyrosine analogue substituted at position 3.

[0046] Moreover, it is obvious that the method of the invention can be widely applied for the modification of tyrosyl-tRNA synthetases in other biological species since the amino acid residues at positions 37 and 195 in the wild-type TyrRS of *E. coli* used here are highly conserved in biological species.

[0047] The translation system utilized here is the wheat germ cell-free translation system, but it is obvious that the unnatural amino acid can be introduced site-specifically by the use of the VC variant made in the invention similarly in the cell-free translation systems of the other eukaryotic organisms and in eukaryotic cells. Also, the protein where all tyrosine residues are replaced by the unnatural amino acid can be made in a prokaryotic translation system by the use of this enzyme. The generation of such a protein also becomes possible by modifying the tyrosyl-tRNA synthetase derived from the other prokaryotic organisms as with the invention. On the other hand, it becomes possible to generate proteins containing the unnatural amino acids site-specifically in the prokaryotic translation systems by modifying the tyrosyl-tRNA synthetase derived from an eukaryotic organism or archaebacteria as with the invention.

[0048] The present invention is aimed to provide the enzyme variants in which the substrate specificity of the tyrosyl-tRNA synthetase is altered. More particularly, the invention is characterized in that the variant of tyrosyl-tRNA synthetase capable of producing the alloprotein in a large quantity containing the unnatural amino acid efficiently and selectively is prepared by modifying the amino acid sequence of the tyrosyl-tRNA synthetase.

[0049] The tyrosyl-tRNA synthetase of the invention is characterized by being capable of selectively, preferentially, and efficiently incorporating a tyrosine analogue substituted at position 3 such as 3-halotyrosine such as 3-iodotyrosine, and 3-hydroxytyrosine known to have various physiological activities as the unnatural amino acid. Also, the tyrosyl-

tRNA synthetase of the invention may be those derived from any organisms such as bacteria such as *E. coli*, yeast, animals and plants, but it is preferable that enzymes with known three dimensional structure and those which are highly conservative as described previously are widely used. Also, even if the tyrosyl-tRNA synthetase is specific for the certain organism, if its variant is not harmful for another organism, it acts in the organism as the tyrosyl-tRNA synthetase for the new amino acid, and thus it can be used as the preferable variant.

5 [0050] As the method to determine the position where the amino acid sequence of the tyrosyl-tRNA synthetase of the invention is modified, it is preferable to determine the position of the amino acid to be modified on the basis of three dimensional structures of the tyrosyl-tRNA synthetase and the enzyme-bound amino acid or AMP complex of the amino acid (aminoacyl-AMP and aminoacyl-AMP analogues), but the method is not limited thereto. On the basis of the case 10 of the tyrosyl-tRNA synthetase exemplified above where the amino acids were at the residues 37 and 195, the position of amino acid to be modified can be determined at the corresponding site in the Rossman fold domain of the other class I aminoacyl-tRNA synthetases or in the anti-parallel  $\beta$ -sheet domain of the class II aminoacyl-tRNA synthetases.

15 [0051] In the invention, it is preferable to choose two or more amino acid residues as the positions to be modified, whereby having been capable of remarkably improving the selectivity for the unnatural amino acid.

20 [0052] When the positions of amino acid to be modified are determined on the basis of the three dimensional structure, the positions of amino acid to be modified can be determined by first analyzing the tertiary structure of aminoacyl-tRNA synthetase of the target amino acid and subsequently analyzing the positions to which the AMP complex of the target amino acid is bound in that tertiary structure.

25 [0053] Also, in the above example, the modified positions included the residues 37 and 195 because the tyrosyl-tRNA synthetase derived from *E. coli* was used. However, this is the case of *E. coli*. When the tyrosyl-tRNA synthetase derived from another organism is used, the amino acids corresponding to the residues 37 and 195 of tyrosyl-tRNA synthetase derived from *E. coli* in the tertiary structure are the positions to be modified. In the invention, "corresponding position in the three dimensional structure" indicates the position of amino acid corresponding to the position which recognizes the amino acid or the AMP complex of the amino acid (aminoacylated AMP and aminoacylated AMP analogues and the like) in the tertiary structure of the aminoacyl-tRNA synthetase.

30 [0054] As the amino acid replacement newly introduced into the tyrosyl-tRNA synthetase, in consideration of hydrophilicity and length for hydrogen bond of the amino acid, for example, in the case such that the binding to the aminoacyl-AMP is not formed at the position, the amino acid can be modified with a hydrophobic amino acid, or the distance can be modulated by altering glutamine to asparagine.

35 [0055] In the above example of the invention, tyrosine at position 37 of the tyrosyl-tRNA synthetase was modified with valine, leucine, isoleucine or alanine, and glutamine at position 195 was modified with alanine, cysteine, serine or asparagine, but the modification is not limited thereto.

40 [0056] The method for producing the variant of tyrosyl-tRNA synthetase of the invention wherein the amino acids at the certain positions are modified with the other amino acids is preferably carried out by gene engineering technology known in the art. For instance, the variant enzymes can be simply produced by amplifying DNA modified with the nucleotide sequence encoding the amino acid to be modified using the primers having the modified nucleotide sequence encoding the position of the target amino acid to be modified, subsequently by combining the amplified DNA fragments to obtain expression vector encoding the variant of aminoacyl-tRNA synthetase in full length, and by expressing the enzyme using host cells such as *E. coli*. The primers used in this method are from about 20 to 70 bases, and preferably from about 20 to 50 bases. It is preferable to use the relatively long primers, for example, those of 20 or more bases because this primer has 1 to 3 base mismatches with the original base sequence before the modification.

45 [0057] Also, the method for producing the variants of the tyrosyl-tRNA synthetase of the invention wherein the amino acids at the certain positions are modified with other amino acids is not limited to the method described above, and can use various gene engineering technologies such as point mutation technology and methods where the modified fragment is introduced by restriction enzymes known in the art.

50 [0058] The invention provides cells transformed using the DNA encoding the variant of tyrosyl-tRNA synthetase described above. Such cells may be either prokaryotic or eukaryotic cells.

55 [0059] Also, when the variant of tyrosyl-tRNA synthetase of the invention expressed in cells is used for protein synthesis in the cells as such, the cells responsible for the purpose can be used.

[0060] The methods known in the art can be employed as the method for the transformation.

[0061] The invention provides the method for producing the proteins containing the unnatural tyrosine analogues substituted at position 3 using the variants of tyrosyl-tRNA synthetase of the invention having the modified amino acid sequences described above. As described above, since the variant of the invention attaches specifically to the unnatural amino acid with 10 folds or more of selectivity, the protein where the target unnatural amino acid is introduced in place of the natural amino acid can be produced by introducing this variant or the gene encoding this variant into the cells followed by the expression thereof.

[0062] Also, the variants of tyrosyl-tRNA synthetase of the invention having the modified amino acid sequences include not only the use in the cell but also the use in vitro (cell-free system).

[0063] Therefore, the invention provides the method for producing the alloproteins efficiently, selectively, in particular site-selectively and in large amounts.

Examples

5

[0064] The present invention is more specifically illustrated below by examples, but it is not limited to these examples at all.

10 Example 1 (Preparation of an expression vector for TyrRS gene)

15

[0065] A gene of tyrosyl-tRNA synthetase (TyrRS) was cloned from wild-type *E. coli*, W3110. Next, a plasmid pETYRS was made by inserting the TyrRS gene into pET26b which is a vector of *E. coli*.

[0066] Using this plasmid containing the TyrRS gene of *E. coli* as the template, a fragment amplified by PCR using the primer (1) and (2) having the following base sequences:

20

Primer (1): 5'-GGAATTCCATATGGCAAGCAGTAACCTGATTAAACAATT

25 GCAAG-3'

Primer (2): 5'-GCCGAAGCTTGTGACTTCCAGCAAATCAGACAGTAATTCT

30 TTTTACCG-3'

was cleaved with NdeI and HindIII, and subsequently ligated at the NdeI-HindIII site of the vector pET26b to make the expression vector of TyrRS, pET-YRS.

35

Example 2 (Determination of positions where the mutations are introduced).

[0067] Since TyrRS is highly conserved in biological species; using the three dimensional structure of TyrRS and tyrosyl-AMP complex in *Bacillus stearothermophilus* already reported (Brick et al., J. Mol. Bio., 208:83- 1988), the residues 37 and 195 were found where the mutations should be introduced to incorporate the tyrosine analogues substituted at position 3.

Example 3 (Introduction of site-specific mutations)

40

[0068] Oligonucleotides including amino acid substitution to be modified were synthesized as respective primers (mutation introducing primers), (3) to (8) by the standard method.

(a) Generation of a single amino acid substitution where one amino acid was replaced

45

[0069] The DNA sequence encoding the single amino acid substitution where one amino acid at position 37 or 195 was replaced was generated using the primers (3) to (8) shown below. The primers (3) and (4) are for the modification at position 37. And the primers (5) to (8) are for the modification at position 195.

50

Primer (3): 5'-AGGATCGAAGCCGCAAGCGAGCGCGATCGGGCCTTGC GCC-3'

55

Primer (4): 5'-AGGATCGAAGCCGCAAGCGAGCGCGATCGGGCCTTGC GCC-3'

[0070] M denotes C or A, and N denotes A, C, G or T.

5 Primer (5): 5'-ACGGTGTGGTGCTGTCTATTGGTGGTTCTGACC-3'

10 Primer (6): 5'-ACGGTGTGGTGCTGGCAATTGGTGGTTCTGACC-3'

15 Primer (7): 5'-ACGGTGTGGTGCTGAACATTGGTGGTTCTGACC-3'

20 Primer (8): 5'-ACGGTGTGGTGCTGTGCATTGGTGGTTCTGACC-3'

(b) Generation of a double amino acid substitution where two amino acids were replaced

[0071] The DIVA sequence encoding the double amino acid substitution was generated from the plasmids encoding respective single amino acid substitutions at position 37 and 195 by an overlap extension method using the primers, and introduced into the NdeI-BanHI site of pET-YRS. The overlap extension method was performed by purifying two fragments amplified using a pair of the primers (1) and (10) and a pair of the primers (9) and (11), and by amplifying by PCR using the primers (1) and (9) therewith.

25 Primer (1): 5'-GGAATTCCATATGGCAAGCAGTAAC TTGATTAAACAATT  
GCAAG-3'

30 Primer (10): 5'-GATCATCTGGTTAACGGAGAAGTGTGTTGCC-3'

35 Primer (9): 5'-TTCTTCGGATCCAACCGAGACTGCCCGCCTTC-3'

40 Primer (11): 5'-GACCTTCCTGCGCGATATTGGCAAAC-3'

45 Example 4 (Expression of wild-type and variant-type TyrRS)

[0072] Each completely mutated DNA fragment obtained from the above process was inserted into an original position of the plasmid pETYRS. *E. coli*/BLR(DE3) was transformed with the plasmid containing the wild-type or a variant TyrRS gene by the transformation according to Hanahan's method (J. Mol. Bio., 166:557-580).

[0073] 50 Transformants having respective plasmids were isolated, and then cultured in LB medium at 37°C for 16 hours. The wild-type and variants TyrRS proteins were purified from crude extracts of *E. coli* as described below. All processes were performed at 4°C.

[0074] 55 The extract was applied on Ni-agarose column (0.5 mL) equilibrated with buffer A (50 mM Tris-10 mM Mg (OAc)<sub>2</sub>, pH 7.9, 0.3 M NaCl, 5 mM mercaptoethanol) containing 10 mM imidazole. The column was washed with the buffer A containing 40 mM imidazole and eluted with the buffer A containing 250 mM imidazole. Fractions with the activity were pooled, dialyzed against buffer B (100 mM Tris-HCl, pH 7.6, 40 mM-KCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT), and then glycerol was added at a final concentration of 50% thereto, which was then stored at -20°C.

[0075] In the above process, the enzyme variants were obtained which were produced by the substitution at position

Tyr 37 with the amino acid selected from the group consisting of alanine, valine, isoleucine and leucine; and the substitution at position Gln 195 with the amino acid selected from the group consisting of serine, cysteine, asparagine and alanine.

[0076] Hereinafter, these are referred to as 37A, 37V, 37I, 37L, 195S, 195C, 195N; and 195A, respectively. Additionally, the enzymes with double mutations were obtained which were produced by the substitutions both at positions Tyr 37 and Gln 195. Hereinafter, these are referred to as AS, VS, IS, LS, AC, VC, IC, LC, AN, VN, IN, LN, AA, VA, IA and LA, respectively. These are collectively shown in Table 1.

[0077] All enzymes used for gene engineering were purchased from Takara, Toyobo and New England Biolabs (MA, USA). The conditions for gene engineering were determined according to advices from the supplier of the enzyme.

10

#### Example 5 (Substrate specificity assay)

[0078] In order to evaluate the substrate specificity of these enzyme variants, the following three types of assay methods were used.

[0079] The method where the aminoacylation reaction is measured by quantifying inorganic phosphate produced by pyrophosphatase-driven hydrolysis of pyrophosphate which is one of reaction products of the aminoacylation has been described by Lloyd et al. (Nucleic Acids Research, 23:2886-2892, 1995). In one method of the invention, this method was simplified and the aminoacylation reaction was measured by the detection of inorganic phosphate using Biomol green (Funakoshi). Each enzyme obtained was incubated with 0.05 ml of the buffer B containing 1 mM ATP (Seikagaku Corporation), 30  $\mu$ M purified crude tyrosyl-tRNA, and 100  $\mu$ M tyrosine or 0.5 mM 3-iodotyrosine at 37°C for one hour. To stop the reaction, 0.1 ml of Bio mol green was added. Absorbance of its supernatant was measured at 630 nm. The results are shown in Tables 2 and 3.

[0080] Table 2 shows the aminoacylation activity of the wild-type enzyme and eight enzymes with a single amino acid mutation by the symbol. Table 3 shows the aminoacylation activity of the wild-type enzyme, 37V, and 16 enzymes with double mutations by the same symbol. It is shown that the activity is high in the order of \*\*\* > \*\* > - in each Table.

[0081] The enzyme variants obtained by the substitution at positions Tyr 37 and Gln 195 with the other amino acids substantially increased the aminoacylation activity for the tyrosine analogue substituted at position 3. At the same time, relatively, the tyrosylation activity of these proteins is decreased. Therefore, the relative ratios of tyrosine analogue to tyrosine of the enzyme variants are higher than that of the wild-type enzyme.

[0082] In the second method, the activity was assayed by Lloyd's method (Nucleic Acids Research, 23:2886-2892, 1995). Each enzyme obtained was incubated with 0.4 ml of the buffer B containing 1 mM ATP (Seikagaku Corporation), 100  $\mu$ M purified crude tyrosyl-tRNA, 200  $\mu$ M tyrosine or 3-iodotyrosine at 37°C for 6 min. The aminoacylation activity was represented as concentrations of released phosphate per min. The result is shown in Table 4.

[0083] Table 4 shows the aminoacylation activity of the enzymes with double mutations. Four enzymes with double mutations combine advantages of both mutations at positions Tyr 37 and Gln 195. For instance, the relative ratio of tyrosine analogue to tyrosine of VC is more than 10, and the ratio is higher due to the further mutation at position Gln 195 than that of 37V, the enzyme with a single mutation. VN, VS, and AC are inferior to VC in the relative ratio of tyrosine analogue to tyrosine, but have the higher relative ratio of tyrosine analogue to tyrosine than the substitutions with a single amino acid.

[0084] From the result shown in Table 4, it can be concluded that the combination of the mutations at position 37 and 195 could dramatically alter the substrate specificity of the enzyme and increase the aminoacylation activity for the tyrosine analogue substituted at position 3. Among the enzyme variants, VC, VN and AC have the most remarkable property and applicability.

[0085] In the third method, the aminoacylation activity for tyrosine or 3-iodotyrosine for tRNA in full length was assayed by the same method as acidic, polyacrylamide electrophoresis by the method of Wolfson et al. (Wolfson, A.D. et al., RNA, 4:1019-1023, 1998). Tyrosine or iodotyrosine at 200  $\mu$ M and the enzyme at 50 nM were incubated at 37°C for 3 min, and subsequently the electrophoresis was carried out. The result of the aminoacylation activity of the enzymes with double mutations are shown on the photo, substitute for a drawing, in Figure 1. In Figure 1, left three lanes are from the wild-type enzyme, middle three lanes are from the enzyme modified at position 37 with valine (37V) and right three lanes are from the enzyme modified at positions 37 and 195 with valine and cysteine (VC), respectively. Three lanes in each group show the case without the addition of the amino acid (w/o aa), the case with the addition of 3-iodotyrosine (200  $\mu$ M) (I-Tyr), and the case with the addition of tyrosine (200  $\mu$ M) (Tyr), respectively from the left.

[0086] As a result, it was found that VC had high specificity for the tyrosine analogue.

55

#### Example 6 (Production of protein with site-specific insertion of unnatural amino acids)

[0087] The protein containing iodotyrosine was generated by the reaction according to a batch method in a wheat germ cell-free protein synthesis system using the resultant variant with double mutations, VC and the amber suppressor

tRNA derived from tyrosine tRNA of *E. coli*.

[0088] Cell-free protein synthesis in the wheat germ system was carried out using Proteios cell-free protein synthesis kit (Toyobo). A target protein is a short chain type of c-H-Ras added His tag at its C-terminus via a spacer sequence (Milburn et al., *Science* 247:939-945, 1990). A plasmid used for the expression of the protein was made by inserting a gene corresponding to a sequence where a thrombin recognition sequence and His tag were added to the C-terminus of the short chain type of c-H-Ras at SpeI-Sall site of a plasmid, pEU3-NII attached in the kit. When the amino acid sequence of the translation product consequently predicted is represented by the single character code for amino acids, it is as follows.

10

MTEYKLVVVGAGGVGKSALTIQLIQNHFVD 30

EYDPTIEDSYRKQVVIDGETCLLDILDTAG 60

15

QEEYSAMRDQYMRTGEGFLCVFAINNTKSF 90

EDIHQYREQIKRVKDSDDVPMVLVGNKCDL 120

20

AARTVESRQAQDLARSYGIPYIETSAKTRG 150

GVEDAFYTLVREIRQHKLRLGSLVPRGSH 180

25

HHHHH 185

[0089] The translation reaction was carried out by further adding iodotyrosine at a final concentration of 0.6, mM in addition to the condition of the kit. The product was electrophoresed using 4-12% NuPAGE Bis-Tris gel (Invitrogen). Using an image analyzer FLA-2000 (Fujifilm, Tokyo, Japan), counts of RI of bands on the gel were quantified, or protein bands stained using Cypro-tangerine protein gel stain (Molecular probes) on the gel were quantified.

[0090] In the translation using mRNA made from a construct where the 32nd codon of this protein was replaced by the amber codon as the template, the production of the protein in full length was dependent on coexistence of both TyrRS and the suppressor tRNA (Figure 2). This proves that the suppressor tRNA used is scarcely aminoacylated by the endogenous aminoacyl-tRNA synthetase in the wheat germ. The synthesized quantity by the amber suppression was 30 to 40% of the control with no amber codon. In order to increase the synthesized quantity of the protein, the reaction by the suppression was carried out for 3 days using a dialysis method. In this case, it was found that a quantity of the production was 0.1 mg or more in terms of the quantity per ml of reaction solution and that the protein could be synthesized sufficiently in quantity to perform various experiments using proteins.

40

Example 7 (Mass spectrometric analysis of the products in cell-free protein synthesis)

[0091] It was examined by mass spectrometric analysis that iodotyrosine was actually inserted at the position corresponding to the amber codon of the peptide chain. The synthesized protein was purified using Ni-NTA magnetic agarose beads (Qiagen) for LC-MS analysis. The purified protein was further purified by electrophoresis using NuPAGE Bis-Tris gel, and a band of 21 kDa corresponding to the target product was cut out from the gel. Digestion using 0.1  $\mu$ g of *Achromobacter* protease I (Lys-C) was carried out in 0.05 M of Tris-HCl (pH 9) containing 0.1% SDS in a gel fragment at 37°C for 12 hours. Peptides produced by the digestion were separated using model 1100 liquid chromatography (Hewlett Packard), connecting DEAE-5PW (1 x 10 mm, Tosoh, Tokyo) and Mightysil C18 (1 x 50 mm, Kanto Chemical, Tokyo) by solvent B with a linear gradient of 2 to 60% over 30 min. The detection was carried out by a diode-array detector. Solvents A and B were an aqueous solution of 0.09% (v/v) trifluoroacetate and 0.075% (v/v) trifluoroacetate-80% (v/v) acetonitrile, respectively. Elutes were analyzed by Finnigan LCQ ion trap mass spectrometer using ESI probe. Tandem MS sequencing was carried out for the fragment corresponding to  $[M + 2H]^{2+}$  of Lys-C fragment containing the 32nd amino acid of Ras.

[0092] Among the fragments obtained by digesting the translation products, the fragment containing the 32nd amino acid from the N-terminus (Ser 17 to Lys 42, referred to as K3 fragment) was focused and analyzed. As a result, the majority of the K3 fragments produced by the amber suppression had the mass corresponding one where one residue of tyrosine was replaced by iodotyrosine, and only minor fragments had the mass not including iodotyrosine. No K3 fragment was detected where the other amino acid was accidentally inserted at the site corresponding to the amber

codon. The ratio of iodotyrosine to tyrosine inserted to the amber codon was examined by absorbance of chromatogram. Consequently, it was found that 95% or more of K3 fragments encompassed iodotyrosine by the amber suppression whereas the ratio of tyrosine insertion at the amber codon was 5% or less. On the other hand, no fragment having the mass corresponding to the case encompassing iodotyrosine in place of tyrosine was detected in the analysis of the 5 translation products from mRNA which did not contain the amber codon. This confirmed that the canonical codon for tyrosine was not contaminated with iodotyrosine (Figure 3).

[0093] Additionally, Phe-Val-Asp-Glu-(iodo-Tyr)-Asp which is a partial sequence of this fragment was confirmed by the tandem mass analysis of the K3 fragment containing iodotyrosine.

10 Industrial Applicability

[0094] The present invention provides novel variants of tyrosyl-tRNA synthetase and novel methods for the modification. According to the method for the modification of the invention, the modifications of selectivity and substrate specificity of tyrosyl-tRNA synthetase can be efficiently carried out by the simple method.

15 [0095] The use of the variant of tyrosyl-tRNA synthetase of the invention enables site-specific insertion of tyrosine analogue substituted at position 3, in addition to natural amino acids, by utilizing aminoacylation reaction upon the preparation of peptide or protein.

[0096] Also, the use of the variant of tyrosyl-tRNA synthetase of the invention makes it possible to produce alloproteins suitable for mass production efficiently and selectively.

20

25

30

35

40

45

50

55

## SEQUENCE LISTING

5

10

&lt;110&gt; Japan Science and Technology Agency

15

&lt;120&gt; TYROSYL-TRNA SYNTHETASE VARIANTS

20

&lt;130&gt; JA909400

25

&lt;150&gt; JP2001-234135

30

&lt;151&gt; 2001-08-01

&lt;160&gt; 13

35

&lt;170&gt; PatentIn Ver. 2.1

40

&lt;210&gt; 1

&lt;211&gt; 424

&lt;212&gt; PRT

&lt;213&gt; Escherichia Coli

45

&lt;400&gt; 1

Met Ala Ser Ser Asn Leu Ile Lys Gln Leu Gln Glu Arg Gly Leu Val

50

1

5

10

15

55

Ala Gln Val Thr Asp Glu Glu Ala Leu Ala Glu Arg Leu Ala Gln Gly

	20	25	30	
5	Pro Ile Ala Leu Tyr Cys Gly Phe Asp Pro Thr Ala Asp Ser Leu His			
	35	40	45	
10	Leu Gly His Leu Val Pro Leu Leu Cys Leu Lys Arg Phe Gin Gin Ala			
15	50	55	60	
	Gly His Lys Pro Val Ala Leu Val Gly Gly Ala Thr Gly Leu Ile Gly			
20	65	70	75	80
	Asp Pro Ser Phe Lys Ala Ala Glu Arg Lys Leu Asn Thr Glu Glu Thr			
25	85	90	95	
	Val Gln Glu Trp Val Asp Lys Ile Arg Lys Gln Val Ala Pro Phe Leu			
30	100	105	110	
35	Asp Phe Asp Cys Gly Glu Asn Ser Ala Ile Ala Ala Asn Asn Tyr Asp			
	115	120	125	
40	Trp Phe Gly Asn Met Asn Val Leu Thr Phe Leu Arg Asp Ile Gly Lys			
45	130	135	140	
	His Phe Ser Val Asn Gln Met Ile Asn Lys Glu Ala Val Lys Gln Arg			
50	145	150	155	160
55	Leu Asn Arg Glu Asp Gln Gly Ile Ser Phe Thr Glu Phe Ser Tyr Asn			
	165	170	175	

1 Leu Leu Gln Gly Tyr Asp Phe Ala Cys Leu Asn Lys Gln Tyr Gly Val  
 5 180 185 190  
 10 Val Leu Gln Ile Gly Gly Ser Asp Gln Trp Gly Asn Ile Thr Ser Gly  
 195 200 205  
 15 Ile Asp Leu Thr Arg Arg Leu His Gln Asn Gln Val Phe Gly Leu Thr  
 210 215 220  
 20 Val Pro Leu Ile Thr Lys Ala Asp Gly Thr Lys Phe Gly Lys Thr Glu  
 225 230 235 240  
 25 Gly Gly Ala Val Trp Leu Asp Pro Lys Lys Thr Ser Pro Tyr Lys Phe  
 245 250 255  
 30 Tyr Gln Phe Trp Ile Asn Thr Ala Asp Ala Asp Val Tyr Arg Phe Leu  
 260 265 270  
 35 Lys Phe Phe Thr Phe Met Ser Ile Glu Glu Ile Asn Ala Leu Glu Glu  
 275 280 285  
 40 Glu Asp Lys Asn Ser Gly Lys Ala Pro Arg Ala Gln Tyr Val Leu Ala  
 290 295 300  
 45 Glu Gln Val Thr Arg Leu Val His Gly Glu Glu Gly Leu Gln Ala Ala  
 305 310 315 320  
 50 Lys Arg Ile Thr Glu Cys Leu Phe Ser Gly Ser Leu Ser Ala Leu Ser  
 325 330 335

5 Glu Ala Asp Phe Glu Gln Leu Ala Gln Asp Gly Val Pro Met Val Glu  
 340 345 350

10 Met Glu Lys Gly Ala Asp Leu Met Gln Ala Leu Val Asp Ser Glu Leu  
355 360 365

15 Gln Pro Ser Arg Gly Gln Ala Arg Lys Thr Ile Ala Ser Asn Ala Ile  
370 375 380

20 Thr Phe Asn Gly Glu Lys Gln Ser Asp Pro Glu Tyr Phe Phe Lys Glu  
 385 390 395 400

25 Glu Asp Arg Leu Phe Gly Arg Phe Thr Leu Leu Arg Arg Gly Lys Lys  
405 410 415

Asn Tyr Cys Leu Ile Cys Trp Lys

40 <210> 2

211 > 44

<212> DNA

45 <213> Artificial Sequence

50 *... <220>*

### <223> Description of Artificial Sequence:Primer

55

5                   <221> primer\_bind  
6                   <222> (1)..(44)  
7  
8  
9  
10                  <400> 2  
11                  ggaatccat aaggcaagca gtaacatgtat taaacaaatgg caag           44  
12  
13  
14  
15                  <210> 3  
16                  <211> 50  
17                  <212> DNA  
18                  <213> Artificial Sequence  
19  
20  
21                  <220>  
22                  <223> Description of Artificial Sequence:Primer  
23  
24  
25                  <220>  
26                  <221> primer\_bind  
27                  <222> (1)..(50)  
28  
29  
30                  <400> 3  
31                  gccgaaggctt gtcgacttttc cagcaaaatca gacagtaattt ctttttacgg           50  
32  
33  
34  
35                  <210> 4  
36                  <211> 40  
37                  <212> DNA  
38                  <213> Artificial Sequence  
39  
40  
41                  <220>

## &lt;223&gt; Description of Artificial Sequence:Primer

5           <220>  
10          <221> primer\_bind  
15          <222> (1)..(40)  
20          <400> 4  
25          aggatcgaag ccgcaagcga gcgcgatcgg gccttgcgcc           40  
30          <210> 5  
35          <211> 40  
40          <212> DNA  
45          <213> Artificial Sequence  
50          <220>  
55          <223> Description of Artificial Sequence:Primer  
60          <220>  
65          <221> primer\_bind  
70          <222> (1)..(40)  
75          <400> 5  
80          aggatcgaag ccgcacmnga gcgcgatcgg gccttgcgcc           40  
85          <210> 6  
90          <211> 33  
95          <212> DNA

5           <213> Artificial Sequence

10           <220>

15           <223> Description of Artificial Sequence:Primer

20           <220>

25           <221> primer\_bind

30           <222> (1)..(33)

35           <400> 6

40           acgg!g!gg! gctgiciattt ggtggttc!g acc           33

45           <210> 7

50           <211> 33

55           <212> DNA

60           <213> Artificial Sequence

65           <220>

70           <223> Description of Artificial Sequence:Primer

75           <220>

80           <221> primer\_bind

85           <222> (1)..(33)

90           <400> 7

95           acgg!g!gg! gctggcaattt ggtggttc!g acc           33

5                   <210> 8  
10                <211> 33  
15                <212> DNA  
20                <213> Artificial Sequence  
25                <220>  
30                <223> Description of Artificial Sequence:Primer  
35                <220>  
40                <221> primer\_bind  
45                <222> (1)..(33)  
50                <400> 8  
55                acgggtgggt gcgtggatccgggggttcgt acc 33  
60                <210> 9  
65                <211> 33  
70                <212> DNA  
75                <213> Artificial Sequence  
80                

---

  
85                <220>  
90                <223> Description of Artificial Sequence:Primer  
95                <220>  
100               <221> primer\_bind  
105               <222> (1)..(33)  
110               <400> 9



5           <222> (1)..(30)

10           <400> 11

15            gatcatcgg ttaacggaga agtggggcc                           30

20

25           <210> 12

30           <211> 26

35           <212> DNA

40           <213> Artificial Sequence

45           <220>

50           <223> Description of Artificial Sequence:Primer

55           <220>

60           <221> primer\_bind

65           <222> (1)..(26)

70

75           <400> 12

80            gaccctccctcg cgcgataatcg gcaaac                           26

85

90           <210> 13

95           <211> 185

100           <212> PRT

105           <213> Artificial Sequence

110           <220>

115           <223> Description of Artificial Sequence:Modification of human protein



Ser Ala Lys Thr Arg Gln Gly Val Glu Asp Ala Phe Tyr Thr Leu Val  
 5 145 150 155 160

Arg Glu Ile Arg Gln His Lys Leu Arg Lys Leu Gly Ser Leu Val Pro  
 10 165 170 175

Arg Gly Ser His His His His His  
 15 180 185

20

### Claims

1. Tyrosyl-tRNA synthetase having modified amino acid sequence capable of incorporating unnatural tyrosine analogue substituted at position 3 more efficiently than original natural tyrosine by modifying the amino acid sequence of the tyrosyl-tRNA synthetase.
2. The tyrosyl-tRNA synthetase according to claim 1, wherein an amino acid of tyrosine analogue substituted at position 3 is 3-halogen substituted tyrosine.
3. The tyrosyl-tRNA synthetase according to claim 2, wherein the halogen, is iodine.
4. The tyrosyl-tRNA synthetase according to any of claims 1 to 3, wherein, on the basis of a tertiary structure of the tyrosyl-tRNA synthetase, two or more amino acids where the enzyme thought to recognize tyrosine or tyrosyl-AMP are modified with other amino acids.
5. The tyrosyl-tRNA synthetase according to claim 4, wherein the modified positions of amino acids are at least positions corresponding in the three dimensional structure to tyrosine at position 37 and glutamine at position 195 of the tyrosyl-tRNA synthetase derived from *Escherichia coli*.
6. The tyrosyl-tRNA synthetase according to claim 5, wherein the position corresponding to tyrosine at position 37 in three dimensional structure of the tyrosyl-tRNA synthetase derived from *Escherichia coli* is valine, leucine, iso-leucine or alanine, and the position corresponding to glutamine at position 195 is alanine, cysteine, serine or asparagine.
7. The tyrosyl-tRNA synthetase according to any of claims, 1 to 6, wherein the tyrosyl-tRNA synthetase is derived from bacteria.
8. The tyrosyl-tRNA synthetase according to claim 7, wherein the bacterium is *Escherichia coli*.
9. A method for modifying an amino acid sequence of tyrosyl-tRNA synthetase, characterized in that an amino acid at a position where the tyrosyl-tRNA synthetase attaches to tyrosine or a tyrosyl-AMP is determined on the basis of three dimensional structure and the amino acid at the position is modified with another amino acid when producing the tyrosyl-tRNA synthetase having the modified amino acid sequence. Whereby unnatural tyrosine analogue substituted at position 3 can be more efficiently incorporated than original natural tyrosine by modifying the amino acid sequence of the tyrosyl-tRNA synthetase.
10. The method for modifying the amino acid sequence of the tyrosyl-tRNA synthetase according to claim 9, wherein

the modified amino acids are two or more.

11. A method for producing a protein containing unnatural tyrosine analogue by a protein production system using the tyrosyl-tRNA synthetase having the modified amino acid sequences according to any of claims 1 to 8.

5

12. The method according to claim 11, wherein the protein production system is an intracellular protein translation system.

10

13. The method according to claim 11, wherein the protein production system is a cell-free translation system.

14. A cell transformed by a gene encoding the tyrosyl-tRNA synthetase having the modified amino acid sequence according to any of claims 1 to 8.

15. The cell according to claim 14, wherein the cell is a prokaryotic cell.

15

16. The cell according to claim 14, wherein the cell is a eukaryotic cell.

20

25

30

35

40

45

50

55

FIG. 1

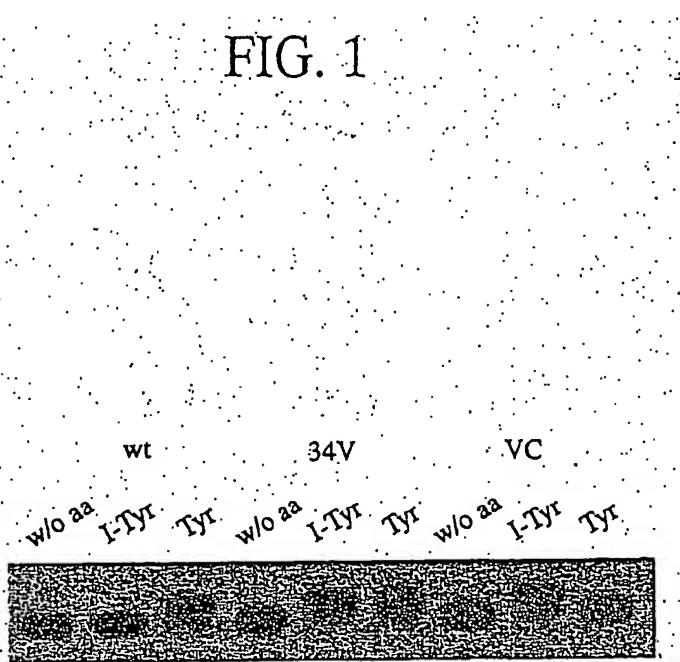
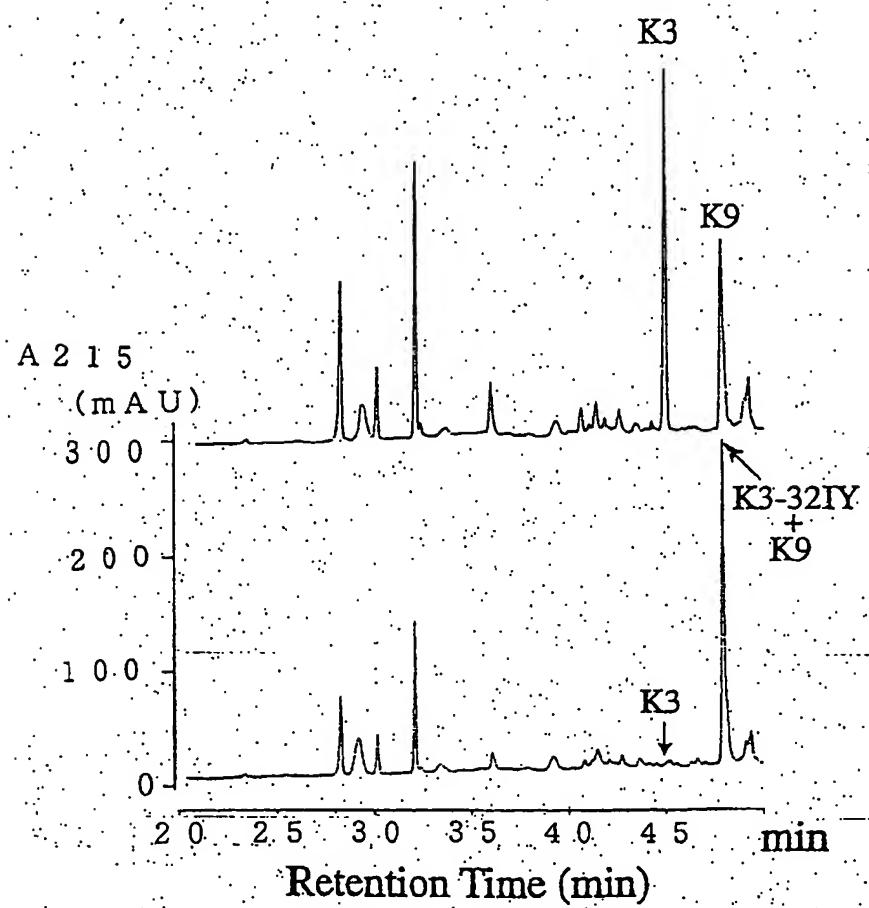


FIG. 2

	mRNA	wt	amb	amb	amb
I-Tyr	+	+	+	+	+
TyrRSvariant	+	+	+	+	-
sup tRNA	+	+	-	-	+
Lane	1	2	3	4	

FIG. 3



INTERNATIONAL SEARCH REPORT		International application No. PCT/JP02/00118												
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> Int.Cl <sup>7</sup> C12N15/52, C12N9/00, C07B61/00, G01N33/48, //G06F17/30, G06F17/50 According to International Patent Classification (IPC) or to both national classification and IPC														
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) Int.Cl <sup>7</sup> C12N15/52, C12N9/00, C07B61/00, G01N33/48, G06F17/30, G06F17/50														
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched														
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) JICST FILE (JOIS), WPI (DIALOG), BIOSIS (DIALOG), MEDLINE (STN), PDB EMBL/DDBJ/Genbank/PIR/Swissprot/Geneseq														
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 2px;">Category*</th> <th style="text-align: left; padding: 2px;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="text-align: left; padding: 2px;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td style="text-align: center; padding: 2px;">P, X</td> <td style="padding: 2px;">OHNO S. et al., Changing the Amino Acid Specificity of Yeast Tyrosyl-tRNA Synthetase by Genetic Engineering, J. Biochem. 01 September, 2001 (01.09.01), Vol.130, No.3, 417-423, Fig.2, table II, table III</td> <td style="text-align: center; padding: 2px;">1-8</td> </tr> <tr> <td style="text-align: center; padding: 2px;">Y</td> <td style="padding: 2px;">HAMANO-TAKAKU F. et al., A Mutant Escherichia coli Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine, J. Biol. Chem. 22 December, 2000 (22.12.00), Vol.275, No.51, pages 40324 to 40328; Fig.3</td> <td style="text-align: center; padding: 2px;">1-8</td> </tr> <tr> <td style="text-align: center; padding: 2px;">Y</td> <td style="padding: 2px;">BAIN J. D. et al., Site-Specific Incorporation of Nonnatural Residues during In Vitro Protein Biosynthesis with Semisynthetic Aminoacyl-tRNAs, Biochemistry (1991), Vol.30, No.22, pages 5411 to 5421, abstract.</td> <td style="text-align: center; padding: 2px;">1-4, 7-8</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	P, X	OHNO S. et al., Changing the Amino Acid Specificity of Yeast Tyrosyl-tRNA Synthetase by Genetic Engineering, J. Biochem. 01 September, 2001 (01.09.01), Vol.130, No.3, 417-423, Fig.2, table II, table III	1-8	Y	HAMANO-TAKAKU F. et al., A Mutant Escherichia coli Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine, J. Biol. Chem. 22 December, 2000 (22.12.00), Vol.275, No.51, pages 40324 to 40328; Fig.3	1-8	Y	BAIN J. D. et al., Site-Specific Incorporation of Nonnatural Residues during In Vitro Protein Biosynthesis with Semisynthetic Aminoacyl-tRNAs, Biochemistry (1991), Vol.30, No.22, pages 5411 to 5421, abstract.	1-4, 7-8
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
P, X	OHNO S. et al., Changing the Amino Acid Specificity of Yeast Tyrosyl-tRNA Synthetase by Genetic Engineering, J. Biochem. 01 September, 2001 (01.09.01), Vol.130, No.3, 417-423, Fig.2, table II, table III	1-8												
Y	HAMANO-TAKAKU F. et al., A Mutant Escherichia coli Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine, J. Biol. Chem. 22 December, 2000 (22.12.00), Vol.275, No.51, pages 40324 to 40328; Fig.3	1-8												
Y	BAIN J. D. et al., Site-Specific Incorporation of Nonnatural Residues during In Vitro Protein Biosynthesis with Semisynthetic Aminoacyl-tRNAs, Biochemistry (1991), Vol.30, No.22, pages 5411 to 5421, abstract.	1-4, 7-8												
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.														
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "B" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family														
Date of the actual completion of the international search 08 April, 2002 (08.04.02)	Date of mailing of the international search report 16 April, 2002 (16.04.02)													
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer													
Facsimile No.	Telephone No.													

INTERNATIONAL SEARCH REPORT		International application No. PCT/JP02/00118
C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BARKER D. G. et al., The tyrosyl-tRNA synthetase from <i>Escherichia coli</i> , FEBS Letters (1982), Vol.150, No.2, pages 419 to 423; Fig. 2	5, 6
Y	WINTER G. et al., The Amino Acid Sequence of the Tyrosyl-tRNA Synthetase from <i>Bacillus Stearothermophilus</i> , Eur. J. Biochem. (1983), Vol.132, No.2, pages 383 to 387; Fig.2	5, 6
Y	BRICK P. et al., Structure of Tyrosyl-tRNA Synthetase Refined at 2.3 Å Resolution, J. Mol. Biol. (1989), Vol.208, No.1, pages 83 to 98 Figs.7, 8	5, 6

Form PCT/ISA/210 (continuation of second sheet) (July 1998)